ASIP FOR MARTINGALES IN 2-SMOOTH BANACH SPACES. APPLICATIONS TO STATIONARY PROCESSES.

CHRISTOPHE CUNY

ABSTRACT. We prove the almost sure invariance principle for martingales with stationary ergodic differences taking values in a separable 2-smooth Banach space (for instance a Hilbert space). A compact law of the iterated logarithm is established in the case of stationary differences of *reverse* martingales. Then, we deduce the almost sure invariance principle for stationary processes under, the Hannan condition, and a compact law of the iterated logarithm for stationary processes arising from non-invertible dynamical systems. Those results for stationary processes are new, even in the real valued case. We also obtain the Marcinkiewicz-Zygmund strong law of large numbers for stationary processes with values in some smooth Banach spaces.

1. INTRODUCTION

Let $(\mathcal{X}, |\cdot|_{\mathcal{X}})$ be a separable Banach space and $\mathcal{X}^*$ be its topological dual. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(X_n)_{n\geq 0}$ be a strictly stationary sequence of $\mathcal{X}$-valued random variables. We are interested in the $\mathbb{P}$-a.s. behaviour of $(S_n/\sqrt{2nL(L(n))})_{n\geq 1}$, where $S_n := X_0 + \cdots + X_{n-1}$ and $L := \max(1, \log)$. We say that $(X_n)_{n\geq 0}$ satisfies the bounded law of the iterated logarithm (bounded LIL or BLIL) if $(S_n/\sqrt{2nL(L(n))})_{n\geq 1}$ is $\mathbb{P}$-a.s. bounded. We say that $(X_n)_{n\geq 0}$ satisfies the compact law of the iterated logarithm (compact LIL or CLIL) if $(S_n/\sqrt{2nL(L(n))})_{n\geq 1}$ is $\mathbb{P}$-a.s. relatively compact.

When $(X_n)_{n\geq 0}$ is a sequence of independent random variables, the bounded and compact LILs are well understood, thanks to a characterization due to Ledoux and Talagrand [23]. When the compact LIL holds, the cluster set of $S_n/\sqrt{2nL(L(n))})_{n\geq 1}$ may be identified thanks to a result of Kuelbs [22]. When X_0 is *pregaussian* (see next section), we have an almost sure invariance principle as well.

For Banach spaces of type 2 (see next section for the definition), the result of Ledoux-Talagrand takes the following particularly simple form.

Theorem 1.1 (Ledoux-Talagrand, [25, Corollary 8.8]). *Let $(X_n)_{n\geq 0}$ be a sequence of iid random variables with values in a Banach space of type 2. Then, $(X_n)_{n\geq 0}$ satisfies the bounded LIL (resp. the compact LIL) if and only if $\mathbb{E}((x^*(X_0))^2) < \infty$ for every $x^* \in \mathcal{X}^*$ (resp. $(((x^*(X_0))^2)_{x^*\in\mathcal{X}^*, |x^*|_{\mathcal{X}^*}\leq 1}$ is uniformly integrable), $\mathbb{E}(|X_0|^2_{\mathcal{X}}/L(L(|X_0|_{\mathcal{X}})) < \infty$ and $\mathbb{E}(X_1) = 0$.*

In particular, a sequence of iid variables $(X_n)_{n\geq 0}$ with values in a Banach space of type 2 satisfies the compact LIL (hence the bounded LIL) as soon as:

$$\mathbb{E}(|X_0|^2_{\mathcal{X}}) < \infty \text{ and } \mathbb{E}(X_0) = 0\,. \tag{1}$$

We are interested here in the case where $(X_n)_{n\geq 0}$ is a general stationary sequence, including the case martingale differences (and of reverse martingale differences). The analogue of the notion of Banach space of type 2 in the case martingale differences is the notion of 2-smooth Banach space (see the next section for the definition). One could wonder whether Theorem

1991 *Mathematics Subject Classification.* Primary: 47A35, 28D05, 37A05; Secondary: 47B38.

Key words and phrases. Banach valued martingales, compact law of the iterated logarithm, strong invariance principle, Hannan's condition.

1.1 is true in this context, or, at least, whether (1) is sufficient for the bounded or compact LIL, when $(X_n)_{n\ge 0}$ is a stationary sequence of martingale differences.

As far as we know, the latter question remained unsolved. Let us however mention some results in that direction. Morrow and Philipp [30] (see also [31] for an improved version) obtained an almost sure invariance principle (see the next section for the definition), hence a compact LIL (with an ad hoc normalization), for sequences of non-necessarily stationary martingale differences taking values in a Hilbert space. Dehling, Denker and Philipp [17] proved a bounded LIL in the same context. When applied to stationary sequences of martingale differences, the above results require higher moments than 2. In [30] and [31] an extra condition on the "conditional variance" is required and a rate in the ASIP is obtained in the finite dimensional case.

In this paper, we prove that condition (1), is sufficient for the compact LIL when $(X_n)_{n\ge 0}$ is a stationary sequence of martingale differences with values in a 2-smooth Banach space. When the sequence is ergodic, the cluster set of $(S_n/\sqrt{2nL(L(n))})_{n\ge 1}$ is identified as well as $\limsup |S_n|_{\mathcal{X}}/\sqrt{nL(L(n))}$. Then, using a result of Berger [2], we obtain an almost sure invariance principle for $(S_n)_{n\ge 1}$. Those results (except for the invariance principle) extend to *reverse* martingale differences.

To prove those results we first obtain integrability properties of the "natural" maximal function arising in that context. This step is crucial not only to prove the results for martingales (and reverse martingales), but also in order to extend the results to general stationary processes under projective conditions. We note that the almost sure invariance principle for Hilbert-valued stationary processes under mixing conditions have been obtained by Merlevède [27] and Dedecker-Merlevède [15]. Their results have different range of application.

We also investigate the Marcinkiewicz-Zygmund strong law of large numbers for stationary processes taking values in a smooth Banach space. The maximal function arising in that other context has been studied by Woyczyński [37], for stationary martingale differences. We investigate the case of stationary processes under projective conditions. The main argument used is the same as the one for the law of the iterated logarithm. The Marcinkiewicz-Zygmund strong laws in smooth Banach spaces have been also investigated by Dedecker-Merlevède [14] for stationary processes satisfying mixing conditions.

In the next section we set our notations and state our results for martingales and then, for stationary processes, including non-adapted processes, functionals of markov chains or iterates of non-invertible dynamical systems. In section 3 we give several examples to which our conditions apply. In section 4 we prove our martingale results and in section 5 we prove our results for stationary processes. When needed, the remarks are proven there as well. Finally we postpone some technical proofs or results to the appendix.

2. Main results

2.1. **Results for stationary (reverse) martingale differences.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We will consider Banach-valued random variables. We refer to Ledoux-Talagrand [25] for our notations and definitions.

Let $(\mathcal{X}, |\cdot|_{\mathcal{X}})$ be a separable Banach space. We endow $\mathcal{X}$ with its Borel σ-algebra. Denote by $L^0(\mathcal{X})$ the space (of classes modulo $\mathbb{P}$) of measurable random variables on Ω taking values in $\mathcal{X}$. We define, for every $p\ge 1$, the usual Bochner spaces L^p and their weak versions, as follows

$$L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) = \{Z\in L^0(\mathcal{X}) \ : \ \mathbb{E}(|Z|_{\mathcal{X}}^p) < \infty\}\,;$$
$$L^{p,\infty}(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X}) = \{Z\in L^0(\mathcal{X}) \ : \ \sup_{t>0} t(\mathbb{P}(|Z|_{\mathcal{X}} > t))^{1/p} < \infty|\}\,.$$

For every $Z \in L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$, write $\|Z\|_{p,\mathcal{X}} := (\mathbb{E}(|Z|^p_{\mathcal{X}}))^{1/p}$ and for every $Z \in L^{p,\infty}(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$, write $\|Z\|_{p,\infty,\mathcal{X}} := \sup_{t>0} t(\mathbb{P}(|Z|_{\mathcal{X}} > t))^{1/p}$.

For the sake of clarity, when they are understood, some of the references to Ω, $\mathcal{F}$, $\mathbb{P}$ or $\mathcal{X}$ may be omitted. Recall that for every $p > 1$ there exists a norm on $L^{p,\infty}(\mathbb{P}, \mathcal{X})$ (see for instance [25], Chapter "Notation"), equivalent to the quasi-norm $\|\cdot\|_{p,\infty}$, that makes $L^{p,\infty}(\mathbb{P}, \mathcal{X})$ a Banach space.

The Banach spaces we will consider are the so-called *smooth* Banach spaces. We refer to Pisier [33] for the definitions and some properties of those spaces. We say that $\mathcal{X}$ is *r-smooth*, for some $1 < r \le 2$, if there exists $L \ge 1$, such that

$$|x+y|^r_{\mathcal{X}} + |x-y|^r_{\mathcal{X}} \le 2(|x|^r_{\mathcal{X}} + L^r|y|^r) \qquad \forall x, y \in \mathcal{X}\,.$$

It is known that when $\mathcal{X}$ is r-smooth, there exists $D \ge 1$, such that for every martingale differences $(d_n)_{1\le n\le N}$, writing $M_N = d_1 + \ldots + d_N$, we have

$$\mathbb{E}(|M_N|^r_{\mathcal{X}}) \le D^r \sum_{n=1}^{N} \mathbb{E}(|d_n|^r_{\mathcal{X}})\,. \tag{2}$$

When needed, we will say that $\mathcal{X}$ is (r, D)-smooth, where D is a constant such that condition (2) be satisfied (notice that this definition is compatible with the definition p. 1680 of [32], see Proposition 2.5 there).

Any L^p space (of $\mathbb{R}$-valued functions) associated with a σ-finite measure is r-smooth for $r = \min(2, p)$. Any Hilbert space is $(2, 1)$-smooth. We say that $\mathcal{X}$ is a Banach space of type r, $1 < r \le 2$, if (2) holds foe every finite set $(d_Nn)_{1\le n\le N}$ of independent variables. In particular, 2-smooth Banach spaces are particular examples of spaces of type 2.

Our goal is to study the law of the iterated logarithm and the Marcinkiewicz-Zygmund strong law of large numbers for the partial sums of a $\mathcal{X}$-valued stationary process. We will start by studying the maximal functions associated with these limit theorems. Let us precise some notations.

Let θ be a measurable measure preserving transformation on Ω. To any $X \in L^0(\Omega, \mathcal{X})$, we associate a stationary process $(X_n)_{n\ge0}$ by setting $X_n = X \circ \theta^n$ (when θ is invertible, we extend that definition to $n \in \mathbb{Z}$). Then, for every $n \ge 1$, write $S_n(X) = \sum_{i=0}^{n-1} X \circ \theta^i$. Let $\mathcal{F}_0 \subset \mathcal{F}$ be a σ-algebra such that $\mathcal{F}_0 \subset \theta^{-1}(\mathcal{F}_0)$ and define a non-decreasing filtration $(\mathcal{F}_n)_{n\ge0}$ by $\mathcal{F}_n := \theta^{-n}(\mathcal{F}_0)$. Define then $\mathbb{E}_n = \mathbb{E}(\cdot|\mathcal{F}_n)$. Let also $\mathcal{F}^0$ be such that $\theta^{-1}(\mathcal{F}^0) \subset \mathcal{F}^0$ (for instance take $\mathcal{F}^0 = \mathcal{F}$) and define a non-increasing filtration $(\mathcal{F}^n)_{n\ge0}$, by $\mathcal{F}^n := \theta^{-n}(\mathcal{F}^0)$. Define then $\mathbb{E}^n = \mathbb{E}(\cdot|\mathcal{F}^n)$.

Let $1 \le p \le 2$. Let $X \in L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$. We consider the following maximal functions

$$\mathcal{M}_p(X) := \sup_{n\ge1} \frac{|\sum_{k=0}^{n-1} X_k|_{\mathcal{X}}}{n^{1/p}}\,, \quad \text{if } 1 \le p < 2, \tag{3}$$

$$\mathcal{M}_2(X) := \sup_{n\ge1} \frac{|\sum_{k=0}^{n-1} X_k|_{\mathcal{X}}}{\sqrt{nL(L(n))}}\,, \tag{4}$$

where $L := \max(\log, 1)$.

The maximal operator $\mathcal{M}_1$ is related to Birkhoff's ergodic theorem, which asserts that for every $X \in L^1(\Omega, \mathcal{X})$, $(\sum_{k=0}^{n-1} X_k)/n)_{n\ge1}$ converges $\mathbb{P}$-a.s. By Hopf's dominated ergodic theorem for real valued stationary processes (see [21, Cor. 2.2 p. 8]), for every $X \in L^1(\Omega, \mathcal{X})$,

$$\|\mathcal{M}_1(X)\|_{1,\infty,\mathcal{X}} \le \|X\|_{1,\mathcal{X}}\,. \tag{5}$$

Now, once we know that (5) holds, by the Banach principle (see [21, Theorem 7.2, p. 64] or Proposition D.1), in order to prove Birkhoff's ergodic theorem, it suffices to prove it on a set of X's dense in L^1 (e.g. the θ invariant elements and the coboundaries). We want to use that strategy to study the Marcinkiewicz-Zygmund strong law of large numbers and versions

of the law of the iterated logarithm. Of course one cannot expect to prove an L^p version of (5) without any further assumption on $(X_n)_{n\geq 0}$. But Woyczyński proved such a result when X is a martingale difference.

We have

Proposition 2.1 (Woyczyński, [37]). *Let $1<p<r\leq 2$ and $D\geq 1$. Let $\mathcal{X}$ be a separable (r,D)-smooth Banach space. There exists $C_{p,r}>0$ such that for every $d\in L^p(\Omega,\mathcal{F}_1,\mathbb{P})$ (resp. $d\in L^p(\Omega,\mathcal{F}^0,\mathbb{P})$), with $\mathbb{E}_0(d)=0$ (resp. $\mathbb{E}^1(d)=0$), we have*

$$\|\mathcal{M}_p(d)\|_{p,\infty,\mathcal{X}}\leq C_{p,r}D^{r/p}\|d\|_{p,\mathcal{X}}\,. \tag{6}$$

Moreover,

$$|S_n(d)|_{\mathcal{X}}/n^{1/p}\to 0 \qquad \mathbb{P}\text{-a.s.} \tag{7}$$

Remarks 2.1a. We do not know whether the proposition is true for p-smooth Banach spaces. Actually, Woyczyński proved that $\mathcal{M}_p(d)$ is in any L^r, $r<p$ and worked with martingale differences (not differences of reverse martingales). But his argument applies to obtain the above proposition. We give the proof of (6) in the appendix, for completeness. The proof of (7) is done in [37]. The argument is very similar to the scalar case. Actually by the Banach principle (see Proposition D.1), using (6), it is enough to show (7) in the scalar case, see for instance the proof of Theorem 2.2.
2.1b. A related result is the Baum-Katz inequality, proved in that context in [14]. As in [14] (and in [37]) the stationarity assumption in Proposition 2.1 may be slightly weakened.

Next, we obtain a similar result for $\mathcal{M}_2$, from which we derive the compact LIL for stationary martingale differences (or reverse martingale differences).

Theorem 2.2. *Let $\mathcal{X}$ be a $(2,D)$-smooth separable Banach space, for some $D\geq 1$. For every $1\leq p<2$, there exists a constant $C_p\geq 1$, such that for every $d\in L^2(\Omega,\mathcal{F}_1,\mathcal{X})$ (resp. every $d\in L^2(\Omega,\mathcal{F}^0,\mathcal{X})$) with $\mathbb{E}_0(d)=0$ (resp. $\mathbb{E}^1(d)=0$), we have*

$$\|\mathcal{M}_2(d)\|_{p,\infty,\mathcal{X}}\leq C_pD\|d\|_{2,\mathcal{X}} \tag{8}$$

In particular, $(d_n)_{n\geq 0}$ satisfies the compact LIL. Moreover, if θ (or the sequence $(d_n)_{n\geq 0}$) is ergodic,

$$\text{and}\qquad \limsup_n \frac{|S_n(d)|_{\mathcal{X}}}{\sqrt{2nL(L(n))}}=\sup_{x^*\in\mathcal{X}^*,|x^*|_{\mathcal{X}^*}\leq 1}\|x^*(d)\|_2\leq\|d\|_{2,\mathcal{X}} \qquad \mathbb{P}\text{-a.s.} \tag{9}$$

Remarks 2.2a. Of course, (8) is equivalent to the fact that, for every $1\leq p<2$, there exists $\tilde{C}_p$, such that $\|\mathcal{M}_2(d)\|_p\leq\tilde{C}_pD\|d\|_{2,\mathcal{X}}$. The explicit estimates in the proof of (8) allows to prove that for every $\varepsilon>0$, $\mathbb{E}((\mathcal{M}_2(d))^2/(L(\mathcal{M}_2(d))^{2+\varepsilon})<\infty$.
2.2b. The maximal inequality (8) implies very directly, the bounded LIL (hence the compact LIL, in the finite dimensional case). Now, by (8) and Proposition D.1, we will see that the proof of the compact LIL reduces to the finite dimensional case. Similarly, using (8), to prove the Hartman-Wintner LIL (for martingales with stationary ergodic differences) in the one-dimensional case, it is enough to prove it when d is bounded.
2.2c. The maximal function $\mathcal{M}_2(X)$ has been already used successfully by Ledoux-Talagrand [23] in the context of iid variables.

When θ is ergodic the cluster set of $\{\frac{S_n(d)}{\sqrt{2nL(L(n))}},\ n\geq 1\}$ may be identified as in [22, Theorem 3.1, II]. We can deduce from Theorem 2.2 an almost sure invariance principle (ASIP). We first give the notations to precise what we mean by an ASIP, in the Banach space setting.

Recall, that we denote by $\mathcal{X}^*$ the topological dual of $\mathcal{X}$. Let $X\in L^2(\Omega,\mathcal{X})$ such that $\mathbb{E}(X)=0$. We define a bounded *symmetric* bilinear operator $\mathcal{K}=\mathcal{K}_X$ from $\mathcal{X}^*\times\mathcal{X}^*$ to $\mathbb{R}$,

by

$$\mathcal{K}(x^*, y^*) = \mathbb{E}(x^*(X)y^*(X)) \qquad \forall x^*, y^* \in \mathcal{X}^* .$$

The operator $\mathcal{K}_X$ is called the *covariance operator* associated to X.

We say that a random variable $W \in L^2(\Omega, \mathcal{X})$ is *gaussian*, if $\mathbb{E}(W) = 0$ and for every $x^* \in \mathcal{X}^*$, $x^*(W)$ has a normal distribution. We say that a random variable $X \in L^2(\Omega, \mathcal{X})$ is *pregaussian*, if $\mathbb{E}(X) = 0$ and there exists a gaussian variable $W \in L^2(\Omega, \mathcal{X})$ with the same covariance operator, i.e. such that $\mathcal{K}_X = \mathcal{K}_W$.

Let $X \in L^2(\Omega, \mathcal{X})$. We say that $(X_n)_{n\geq 0}$ satisfies the almost sure invariance principle (ASIP) if, extending our probability space if necessary, there exists a sequence $(W_n)_{n\geq 0}$ of iid gaussian variables, such that

$$|S_n(X) - (W_0 + \cdots + W_{n-1})|_{\mathcal{X}} = o(\sqrt{nL(L(n))}) \qquad \mathbb{P}\text{-a.s.}$$

We shall say that $(X_n)_{n\geq 0}$ satisfies the ASIP of covariance $\mathcal{K}$, when $\mathcal{K} = \mathcal{K}_{W_0}$ is identified.

We now recall an important result of Berger on the ASIP for martingale differences.

Proposition 2.3 (Berger, [2])**.** *Let $\mathcal{X}$ be a separable Banach space. Assume that θ is ergodic. Let $d \in L^2(\Omega, \mathcal{F}_1, \mathcal{X})$, with $\mathbb{E}_0(d) = 0$. Assume that d is pregausssian and that $(d_n)_{n\geq 0}$ satisfies the CLIL. Then, $(d_n)_{n\geq 0}$ satisfies the ASIP of covariance $\mathcal{K}_d$.*

By [25, Proposition 9.24], on any Banach space $\mathcal{X}$ of type 2 (in particular, on any 2-smooth Banach space), every $X \in L^2(\Omega, \mathcal{X})$ with $\mathbb{E}(X) = 0$ is pregaussian. Hence, Berger's result applies as soon as the CLIL is satisfied and we deduce:

Corollary 2.4. *Let $\mathcal{X}$ be a 2-smooth separable Banach space. Assume that θ is ergodic. For every $d \in L^2(\Omega, \mathcal{F}_1, \mathcal{X})$, with $\mathbb{E}_0(d) = 0$, $(d_n)_{n\geq 0}$ satisfies the ASIP of covariance $\mathcal{K}_d$.*

Remark 2.4. Assume that $\dim \mathcal{X} = 1$ and that θ is ergodic. It follows from Corollary 2.5 of [7] that for $d \in L^2(\Omega, \mathcal{F}^0, \mathcal{X})$ such that $\mathbb{E}^1(d) = 0$, $(d_n)_{n\geq 0}$ satisfies the ASIP. We do not know whether the ASIP holds when $\dim \mathcal{X} \geq 2$. The proof of Proposition 2.3 given in [2] does not seem to pass to *reverse* martingale differences. One possibility of proof could be to apply proposition 3.1 of [12], using (8) to verify assumption 1 there.

2.2. **Results for non necessarily adapted stationary processes.** We assume here that θ is invertible and bi-measurable, in which case we extend our filtration to $(\mathcal{F}_n)_{n\in\mathbb{Z}}$. Then, we write $\mathcal{F}_{-\infty} := \cap_{n\in\mathbb{Z}}\mathcal{F}_n$, $\mathcal{F}_\infty := \vee_{n\in\mathbb{Z}}\mathcal{F}_n$, and for every $n \in \overline{\mathbb{Z}}$, $\mathbb{E}_n(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_n)$ and $P_n := \mathbb{E}_n - \mathbb{E}_{n-1}$. We say that a random variable $X \in L^1(\Omega, \mathcal{X})$ is *regular* if $\mathbb{E}_{-\infty}(X) = 0$ and $X - \mathbb{E}_\infty(X) = 0$.

Theorem 2.5. *Assume that θ is invertible and bi-measurable. Let $1 < p < r \leq 2$ and $D > 0$. Let $\mathcal{X}$ be a (r, D)-smooth separable Banach space and $X \in L^p(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{X})$ be a regular variable. Assume moreover that*

$$\|X\|_{H_p} := \sum_{n\in\mathbb{Z}} \|P_n X\|_{p,\mathcal{X}} < \infty . \tag{10}$$

Then, there exist (a universal) $C_{p,r} > 0$, such that

$$\|\mathcal{M}_p(X)\|_{p,\infty} \leq C_{p,r} D^{r/p} \|X\|_{H_p} . \tag{11}$$

Moreover

$$|S_n(X)|_{\mathcal{X}}/n^{1/p} \to 0 \qquad \mathbb{P}\text{-}a.s.$$

Remark 2.5a. If we assume moreover $\mathcal{X}$ to be a Hilbert space, say $\mathcal{H}$, then condition (10) holds as soon as $\sum_{n\geq 1} \frac{\|\mathbb{E}_{-n}(X)\|_{p,\mathcal{H}}}{\sqrt{n}} < \infty$ and $\sum_{n\geq 1} \frac{\|X - \mathbb{E}_n(X)\|_{p,\mathcal{H}}}{\sqrt{n}} < \infty$.
2.5b. Theorem 2.5 improves Corollary 1 of [38] and part of Corollary 3.1 of [6].

Theorem 2.6. *Assume that θ is invertible and bi-measurable. Let $\mathcal{X}$ be a $(2,D)$-smooth separable Banach space, for some $D\geq 1$. Let $X\in L^2(\Omega,\mathcal{F},\mathbb{P},\mathcal{X})$ be a regular random variable. Assume moreover that*

$$\|X\|_{H_2}:=\sum_{n\in\mathbb{Z}}\|P_nX\|_{2,\mathcal{X}}<\infty\,. \tag{12}$$

Then, for every $1\leq p<2$, there exist (a universal) $C_p>0$, such that

$$\|\mathcal{M}_2(X)\|_{p,\infty}\leq C_pD\|X\|_{H_2}\,. \tag{13}$$

The series $d=\sum_{n\in\mathbb{Z}}P_1(X\circ\theta^n)$ converges in $L^2(\Omega,\mathcal{F}_1,\mathcal{X})$ and $\mathbb{E}_0(d)=0$. Moreover, wrting $M_n:=\sum_{k=0}^{n-1}d_k$, we have

$$|S_n-M_n|_{\mathcal{X}}=o(\sqrt{nL(L(n))})\qquad \mathbb{P}\text{-a.s.} \tag{14}$$

Remarks 2.6a. It follows from (14) that the conclusion of Theorem 2.2 holds for $(X_n)_{n\geq 0}$. In particular $(X_n)_{n\geq 0}$ satisfies the CLIL and the ASIP of covariance $\mathcal{K}_d$, where, for every $x^*,y^*\in\mathcal{X}^*$, $\mathcal{K}_d(x^*,y^*)$, $\mathcal{K}_d=\sum_{n\in\mathbb{Z}}\mathbb{E}(x^*(X_n)y^*(X))$, the series being abslutely convergent. Moreover, since, $\|d\|_{2,\mathcal{X}}\leq\|X\|_{H_2}$,

$$\limsup_n\frac{|S_n(X)|_{\mathcal{X}}}{\sqrt{2nL(L(n))}}\leq\|X\|_{H_2}\quad\mathbb{P}\text{-a.s.}$$

2.6b. Notice that on a Hilbert space $\mathcal{H}$ (see the appendix for the proof of the remark), condition (12) holds as soon as

$$\sum_{n\geq 1}\frac{\|\mathbb{E}_{-n}(X)\|_{2,\mathcal{H}}}{\sqrt{n}}<\infty\ \text{ and }\ \sum_{n\geq 1}\frac{\|X-\mathbb{E}_n(X)\|_{2,\mathcal{H}}}{\sqrt{n}}<\infty\,. \tag{15}$$

Moreover, (15) implies that X is regular.

2.6c. If X is not regular, write $X=X-\mathbb{E}_\infty(X)+Y+\mathbb{E}_{-\infty}(X)$, where $Y=\mathbb{E}_\infty(X)-\mathbb{E}_{-\infty}(X)$ is regular and satisfies $\|Y\|_{H_2}=\|X\|_{H_2}$. Then, the conclusion of Theorem 2.6 holds when the regularity condition is replaced with the following: $|(S_n(\mathbb{E}_{-\infty}(X)))|_{\mathcal{X}}=o(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s. and $|S_n(X-\mathbb{E}_\infty(X))|_{\mathcal{X}}=o(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s. Now, it follows from [6, Theorem 4.7] that a sufficient condition for the latter is $\sum_n\frac{\|\mathbb{E}_{-\infty}(S_n)\|_2}{n^{3/2}(L(L(n)))^{1/2}}<\infty$ and $\sum_n\frac{\|S_n-\mathbb{E}_\infty(S_n)\|_2}{n^{3/2}(L(L(n)))^{1/2}}<\infty$.

2.6d. Theorem 2.6 improves Theorem 2 of [38], Theorem 2.1 of [24] (for $p=2$) and Corollary 5.3 of [5], where the results do not apply to infinite dimensional Banach spaces.

2.6e. By a standard argument (see for instance Proposition 2.1 of [1]), it follows from (13), (8) and (14) that $|\mathbb{E}_0(S_n-M_n)|_{\mathcal{X}}=o(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s. and $|\mathbb{E}_n(S_n-M_n)|_{\mathcal{X}}=o(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s. Since $\mathbb{E}_0(M_n)=0$ and $\mathbb{E}_n(M_n)=M_n$ we deduce that $|\mathbb{E}_0(S_n)|_{\mathcal{X}}=o(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s. and $|S_n-\mathbb{E}_n(S_n)|_{\mathcal{X}}=o(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s. On the other hand, one can adapt the proof of Theorem 1 of [36] to prove that for every positive φ with $\varphi(n)/L(L(n)\to 0$, there exists a (linear) process $(X_n)_{n\in\mathbb{Z}}$ satisfying (12) and such that $\limsup|\mathbb{E}_0(S_n)|_{\mathcal{X}}/\sqrt{n\varphi(n)})=+\infty$ $\mathbb{P}$-a.s.

The situation considered in this paragraph includes the case of stationary (ergodic) Markov chains. Let P be a transition probability on a measurable space $(\mathbb{S},\mathcal{S})$ admitting an invariant probability m. Let $(\Omega,\mathcal{F},(\mathcal{F}_n)_{n\in\mathbb{Z}},\mathbb{P},(W_n)_{n\in\mathbb{Z}})$ be the canonical Markov chain associated with P, i.e. $\Omega=\mathcal{S}^{\mathbb{Z}}$, $\mathcal{F}=\mathcal{S}^{\otimes\mathbb{Z}}$, $(W_n)_{n\in\mathbb{Z}}$ the coordinates, $\mathcal{F}_n=\sigma\{\ldots,W_{n-1},W_n\}$, $\mathbb{P}\circ W_0^{-1}=m$ and $\mathbb{P}(W_{n+1}\in A|\mathcal{F}_n)=P(W_n,A)$. Finally, denote by θ the shift on Ω.

Recall that P induces an operator on $L^2(\mathbb{S},m)$ that we still denote by P. If $\mathcal{H}$ is a Hilbert space, we denote by $\mathbf{P}$ the analogous operator on $L^2(\mathbb{S},m,\mathcal{H})$.

Theorem 2.6 applies to that setting with $X=f(W_0)$, where $f\in L^2(\mathbb{S},\mathcal{H})$. Using Remark 2.6b, it suffices to check (15). In that situation, the process is adapted, hence the second part

of condition (15) is automatically satisfied while the first part reads as follows

$$\sum_{n\geq 1}\frac{\|\mathbf{P}^n f\|_{2,\mathcal{H}}}{\sqrt{n}}<\infty\,. \tag{16}$$

If P is normal, i.e. $PP^*=P^*P$, then $\mathbf{P}$ is normal too. In that case, it can be proved that the conclusion of Theorem 2.6 holds for $(X_n=f(W_n))_{n\geq 0}$ as soon as

$$\sum_{n\geq 1}\|\mathbf{P}^n f\|_{2,\mathcal{H}}^2<\infty\,. \tag{17}$$

Condition (17) is clearly weaker than (16).

The sufficiency of (17) is proved in [4], using some arguments of Jiang-Wu [20], see their Theorem 2.1 and their remark 2.2.

2.3. **Results for non-invertible dynamical systems.** Here, we assume that θ is non-invertible. Let us write $\mathcal{F}^n=\theta^{-n}(\mathcal{F})$, for every $n\geq 0$. Denote $\mathcal{F}^\infty=\cap_{n\geq 0}\mathcal{F}^n$.

In this case there exists a Markov operator K, known as the Perron-Frobenius operator, defined by

$$\int_\Omega X\,(Y\circ\theta)\,d\mathbb{P}=\int_\Omega (KX)\,Y\,d\mathbb{P}\qquad \forall X,Y\in L^2(\Omega,\mathcal{F},\mathbb{P})\,. \tag{18}$$

Then, we have for every $X\in L^1(\Omega,\mathcal{F}^0,\mathbb{P})$,

$$\mathbb{E}^n(X)=(K^nX)\circ\theta^n\,. \tag{19}$$

If $\mathcal{H}$ is a Hilbert space, we extend K to $L^2(\Omega,\mathcal{F},\mathbb{P},\mathcal{H})$, in a way similar to (18). We denote by $\mathbf{K}$ the obtained operator.

Theorem 2.7. *Let $(\Omega,\mathcal{F},\mathbb{P},\theta)$ be a non-invertible dynamical system. Let $X\in L^2(\Omega,\mathcal{H})$ be such that*

$$\sum_{n\geq 0}\frac{\|\mathbf{K}^n X\|_{2,\mathcal{H}}}{\sqrt{n}}<\infty\,. \tag{20}$$

Then, for every $1<p<2$, there exists C_p such that

$$\|\mathcal{M}_2(X)\|_{p,\mathcal{H}}\leq C_p\sum_{n\geq 0}\frac{\|\mathbf{K}^n X\|_{2,\mathcal{H}}}{\sqrt{n}}\,.$$

Moreover, there exists $d\in L^2(\Omega,\mathcal{F},\mathbb{P},\mathcal{H})$ with $\mathbb{E}^1(d)=0$, such that, writing $M_n:=\sum_{k=0}^{n-1}d_k$, we have

$$|S_n-M_n|_{\mathcal{H}}=o(\sqrt{nL(L(n))})\qquad \mathbb{P}\text{-}a.s. \tag{21}$$

Remarks 2.7a. It follows from (21) that $(X_n)_{n\geq 0}$ satisfies the CLIL, but we do not know whether it satisfies the ASIP in general, except when $\mathcal{H}$ has dimension one (see Remark 2.2b).
2.7b. Of course, as in the previous paragraph, we have also the CLIL and even the ASIP for the Markov chain induced by $\mathbf{K}$, when (20) is realized.

3. Applications

Now, we give several applications of the previous results. We do not intend to give all possible examples where our conditions apply, but we try to provide examples illustrating the different situations we have considered.

For instance, our results on the Marcinkiewicz-Zygmund strong laws (and on the LIL) can be used (in the one-dimensional case) to obtain almost-sure invariance principles as in [38] (see also [10]).

We start with a one-dimensional situation.

3.1. **ϕ-mixing sequences.** Let us recall the definition of the ϕ-mixing coefficients.

Definition 1. *For any integrable random variable X, let us write $X^{(0)} = X - \mathbb{E}(X)$. For any random variable Y with values in $\mathbb{R}$ and any σ-algebra $\mathcal{F}$, let*

$$\phi(\mathcal{F}, Y) = \sup_{x\in\mathbb{R}} \Big\| \mathbb{E}\Big((\mathbf{1}_{Y\le x})^{(0)} \Big| \mathcal{F}\Big)^{(0)} \Big\|_\infty .$$

For a sequence $\mathbf{Y} = (Y_i)_{i\in\mathbb{Z}}$, where $Y_i = Y_0 \circ \theta^i$ and Y_0 is an $\mathcal{F}_0$-measurable and real-valued random variable, let

$$\phi_{\mathbf{Y}}(n) = \sup_{i\ge n} \phi(\mathcal{F}_0, Y_i) .$$

We need also the following technical definition.

Definition 2. *If μ is a probability measure on $\mathbb{R}$ and $p \in]1,\infty)$, $M \in (0,\infty)$, let $\mathrm{Mon}_p(M,\mu)$ denote the set of functions $f : \mathbb{R} \to \mathbb{R}$ which are monotonic on some interval and null elsewhere and such that $\mu(|f|^p) \le M^p$. Let $\mathrm{Mon}_p^c(M,\mu)$ be the closure in $\mathbb{L}^p(\mu)$ of the set of functions which can be written as $\sum_{\ell=1}^L a_\ell f_\ell$, where $\sum_{\ell=1}^L |a_\ell| \le 1$ and $f_\ell \in \mathrm{Mon}_p(M,\mu)$.*

Theorem 3.1. *Let $X_i = f(Y_i) - \mathbb{E}(f(Y_i))$, where $Y_i = Y_0 \circ \theta^i$ and Y_0 is an $\mathcal{F}_0$-measurable random variable. Let P_{Y_0} be the distribution of Y_0 and $p \in]1,\infty]$. Assume that f belongs to $\mathrm{Mon}_p^c(M, P_{Y_0})$ (with $P_{Y_0} = \mathbb{P}\circ Y_0^{-1}$) for some $M > 0$, if $2 \le p < \infty$ and that f has bounded variation if $p = \infty$. Assume moreover that*

$$\sum_{k\ge1} \frac{\phi_{\mathbf{Y}}^{(p-1)/p}(k)}{k^{1/2}} < \infty. \tag{22}$$

Then, if $1 < p < 2$, $(X_n)_{n\in\mathbb{Z}}$ satisfies the conclusion of Theorem 2.5 and if $p \ge 2$, $(X_n)_{n\in\mathbb{Z}}$ satisfies the conclusion of Theorem 2.6.

Remark. When $p = 2$, Dedecker-Merlevède-Gouëzel [11] proved that $\sum_{n\ge1} \mathbb{P}(\max_{1\le k\le 2^n} |S_k| > C2^{n/2}(L(n))^{1/2}) < \infty$ (which is stronger than (13)) under the condition $\sum_{k\ge1} k^{1/\sqrt{3}-1/2}\phi_{\mathbf{Y}}^{(1/2}(k) < \infty$.

Proof. Assume first that $1 < p < \infty$. Since $f \in \mathrm{Mon}_p(M, P_{Y_0})$, there exists a sequence of functions

$$f_L = \sum_{k=1}^L a_{k,L} f_{k,L} ,$$

such that for every $L \ge 1$, $\sum_{k=1}^L |a_{k,L}| \le 1$, for every $1 \le k \le L$, $f_{k,L}$ is monotonic on some intervall and null elsewhere, and $\|f_{k,L}(Y_0)\|_p \le M$ and, finally $(f_L)_{L\ge1}$ converges in $L^p(P_{Y_0})$ to f. Hence,

$$\|\mathbb{E}_0(f(Y_n)) - \mathbb{E}(f(Y_n))\|_p = \lim_{L\to\infty} \|\mathbb{E}_0(f_L(Y_n)) - \mathbb{E}(f_L(Y_n))\|_p$$
$$\le \liminf_{L\to\infty} \sum_{k=1}^L |a_{k,L}| \|\mathbb{E}_0(f_{k,L}(Y_n)) - \mathbb{E}(f_{k,L}(Y_n))\|_p \le C_p M \phi_{\mathbf{Y}}^{(p-1)/p}(n) ,$$

where we used Lemma 5.2 of [12] for the last estimate.

To conclude in that case, we notice first that we are in the adapted case, and that Theorem 2.5 applies, when $1 < p < 2$, by Remark 2.5a and Theorem 2.6 applies when $p \ge 2$, by remark 2.6b.

Assume that $p = \infty$ and that f has bounded variation. Hence f is the difference of two monotonic functions, to which we apply Lemma 5.2 of [12] with $p = \infty$. Then, we conclude as above. $\square$.

3.2. **Linear processes.** Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space and θ be an ergodic invertible and bi-measurable transformation on Ω. Let $\mathcal{X}$ be a separable 2-smooth Banach space. Let $\xi\in L^2(\Omega,\mathcal{F}_0,\mathbb{P},\mathcal{X})$ such that $\mathbb{E}(\xi|\mathcal{F}_{-1})=0$ and define $\xi_i=\xi\circ\theta^i$, $i\in\mathbb{Z}$.

Let $(A_i)_{i\in\mathbb{Z}}$ be a (non stationary) sequence of random variables with values in $L^\infty(\Omega,\mathcal{F}_{i-1},\mathbf{B}(\mathcal{X}))$, where $\mathbf{B}(\mathcal{X})$ stands for the Banach space of bounded (linear) operators on $\mathcal{X}$. It follows from (2) that, if $\sum_{i\in\mathbb{Z}}\|A_i\|^2_{\infty,\mathbf{B}(\mathcal{X})}<\infty$, the process

$$X_n := \sum_{i\in\mathbb{Z}} A_i\xi_{n-i}\,,\quad n\in\mathbb{Z}$$

is well-defined, in $L^2(\Omega,\mathcal{X})$. We have

Corollary 3.2. *Let (X_n) be the linear process above. Assume moreover that*

$$\sum_{i\in\mathbb{Z}}\|A_i\|_{\infty,\mathbf{B}(\mathcal{X})}<\infty\,.$$

Then X_0 satisfies (12) *and the conclusion of Theorem 2.6 holds.*

3.3. **Functions of linear processes.** Let $(\xi_n)_{n\in\mathbb{Z}}$ be a sequence of iid *real* random variables in $L^2(\Omega,\mathcal{F},\mathbb{P})$. Let $(a_n)_{n\in\mathbb{Z}}$ be in ℓ^1. We consider a linear process defined by

$$Y_n := \sum_{k\in\mathbb{Z}} a_k\xi_{n-k}\qquad \forall n\in\mathbb{Z}\,.$$

For every $n\in\mathbb{Z}$, write $\mathcal{F}_n=\sigma\{\dots,\xi_{n-1},\xi_n\}$.

We denote by Λ the classes of non-decreasing continuous bounded functions on $[0,+\infty[$, such that $\varphi(0)=0$, and satisfying one of the following

$$\varphi^2 \text{ is concave}\quad ;$$
$$\varphi(x)=\min(1,x^\alpha)\qquad \forall x\ge 0\,,\ \text{for some } 0<\alpha\le 1\,.$$

Let $r\ge 1$. Let f be a real valued function such that

$$|f(x)-f(y)|\le\varphi(|x-y|)(1+|x|^r+|y|^r)\qquad \forall x,y\in\mathbb{R}\,. \tag{23}$$

Our functions are unbounded and around a (large) point $x\in\mathbb{R}$ the continuity of f is controlled by φ with an authorized "weight" of $|x|^r$.

We want to study the process $(X_n)_{n\in\mathbb{Z}}$ given by

$$X_n := f(Y_n)-\mathbb{E}(f(Y_n))\qquad \forall n\in\mathbb{Z}\,.$$

Corollary 3.3. *Let $\varphi\in\Lambda$ and $r\ge 1$. Let $\xi_0\in L^{2r}(\Omega,\mathcal{F},\mathbb{P})$ and f satisfy* (23)*. Let $(a_n)_{n\in\mathbb{Z}}\in\ell^1$. Consider the process $(X_n)_{n\ge 0}$ above. If*

$$\sum_{n\ge 1}\varphi(|a_n|)<\infty\qquad or\qquad \sum_{n\ge 1}\frac{\varphi(\sum_{k\ge n}|a_k|)}{\sqrt{n}}<\infty\,,$$

then $(X_n)_{n\ge 0}$ satisfies the conclusion of Theorem 2.6.

We give the proof in the appendix.

3.4. **A non-adapted example.** We now consider an example of non-adapted processes for which new ASIP with rates have been obtained very recently, see [16] and the references therein.

Let $d\ge 2$ and θ be an ergodic automorphism of the d-dimensional torus $\Omega=\Omega_d=\mathbb{R}^d/\mathbb{Z}^d$. Denote by $\mathcal{F}$ the Borel σ-algebra of Ω and take $\mathbb{P}$ to be the Lebesgue measure on Ω.

For every $\mathbf{k}=(k_1,\dots,k_d)\in\mathbb{Z}^d$, write $|\mathbf{k}|:=\max_{1\le i\le d}|k_i|$. If $\mathcal{H}$ is a Hilbert space and if $f\in L^2(\Omega,\mathcal{F},\mathbb{P},\mathcal{H})$, we denote by $(c_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^d}=(c_{\mathbf{k},\mathcal{H}})_{\mathbf{k}\in\mathbb{Z}^d}$ its Fourier coefficients, i.e. $c_{\mathbf{k},\mathcal{H}}=\int_{[0,1]^d}f(x)\mathrm{e}^{-2i\pi\langle x,\mathbf{k}\rangle}\mathbb{P}(dx)$, for every $\mathbf{k}\in\mathbb{Z}^f$, where $\langle\cdot,\cdot\rangle$ stands for the inner product on $\mathbb{R}^d$.

Corollary 3.4. *Let $\mathcal{H}$ be a Hilbert space and $f \in L^2(\Omega, \mathcal{H})$. Assume that there exists $\beta > 2$ and $C > 0$ such that*

$$\sum_{|\mathbf{k}|\geq m} |c_{\mathbf{k}}|^2_{\mathcal{H}} \leq \frac{C}{L(m)(L(L(m)))^\beta} \qquad \forall m \geq 1.$$

Then, $(f\circ\theta^n)_{n\geq 0}$ satisfies the ASIP with covariance operator given by $\mathcal{K}(x,y) := \sum_{m\in\mathbb{Z}} \mathbb{E}(\langle x, f\rangle_{\mathcal{H}} \langle y, f\circ \theta^n\rangle_{\mathcal{H}})$, for every $x, y \in \mathcal{H}$.

Remark 3.4. Dedecker-Merlevède-Pène [16, Theorem 2.1] obtained the ASIP when $\mathcal{H} = \mathbf{R}^m$ and their proof requires $\beta > 4$. When $m = 1$, rates in the ASIP are also provided in [16].
Proof. It follows from the proof of Propositions 4.2 and 4.3 of [16] (notice that the proofs work in the Hilbert space setting) that there exists a filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ (defined at the beginning of paragraph 3 of [16]) such that $\mathcal{F}_n = \theta^{-n}\mathcal{F}_0$ and

$$\|\mathbb{E}_{-n}(f)\|_{2,\mathcal{H}} = O(\frac{1}{\sqrt{nL(n)^\beta}}) \qquad \text{and} \qquad \|\mathbb{E}_n(f) - f\|_{2,\mathcal{H}} = O(\frac{1}{\sqrt{nL(n)^\beta}})\,.$$

Then, the result follows from Remark 2.6b.

4. Proof of the results for Banach-valued martingales

Proof of Theorem 2.2. Let us prove (8). We start with the case $d \in L^2(\Omega, \mathcal{F}_1, \mathbb{P})$ and $\mathbb{E}_0(d) = 0$.

When $d \in L^2(\Omega, \mathcal{F}^0, \mathbb{P})$ and $\mathbb{E}^1(d) = 0$, the proof is the same, with the obvious changes, noticing that for every $n \geq 1$, $(S_n(d) - S_{n-k}(d))_{0\leq k\leq n}$ is a $(\mathcal{F}^{n-k})_{0\leq k\leq n}$-martingale, that $\max_{1\leq k\leq n} |S_k(d)|_{\mathcal{X}} \leq 2\max_{1\leq k\leq n} |S_n(d) - S_{n-k}(d)|_{\mathcal{X}}$ and that the martingale property is only used on blocks.

Clearly, by homogeneity, it suffices to prove the result when $\|d\|_{2,\mathcal{X}} = 1$. Let $\lambda > 0$ and $1 \leq p < 2$. Let us prove that there exists $C_p \geq 1$, independant of λ such that

$$\lambda^p\, \mathbb{P}(M^* > \lambda) \leq D^p C_p^p\,, \tag{24}$$

where

$$M^* = M^*(d) := \sup_{s\geq 0} \frac{\max_{1\leq k\leq 2^s} |M_k|_{\mathcal{X}}}{2^{s/2}(L(s))^{1/2}}\,.$$

Since clearly $\mathcal{M}_2(d) \leq CM^*$, this will imply the desired result. Since (24) is clear for $0 < \lambda < D$, we shall assume that $\lambda \geq D$.

Let $S \geq 1$ be an integer, fixed for the moment.

We have, using Doob's maximal inequality for the submartingale $(|S_n(d)|_{\mathcal{X}})_{n\geq 1}$, and (2)

$$\mathbb{P}(\sup_{1\leq s\leq S} \frac{\max_{1\leq k\leq 2^s} |S_k(d)|_X}{2^{s/2}(L(s))^{1/2}} > \lambda) \leq \frac{1}{\lambda^2}\sum_{s=1}^{S} \frac{\mathbb{E}(\max_{1\leq k\leq 2^s} |S_k(d)|^2_{\mathcal{X}})}{2^s L(s)}$$

$$\leq \frac{2}{\lambda^2}\sum_{s=1}^{S} \frac{\mathbb{E}(|S_{2^s}(d)|^2_{\mathcal{X}})}{2^s L(s)} \leq \frac{CD^2 S}{\lambda^2 L(S)}\,, \tag{25}$$

for some positive constant C.

We make use of truncations. Let $\alpha > 0$ be fixed for the moment. For every $s \geq 1$, $k \geq 1$ define

$$e_k^{(s)} := d_k \mathbf{1}_{\{|d_k|_{\mathcal{X}} \leq \alpha\lambda 2^{s/2}/(L(s))^{1/2}\}} \quad ; \quad d_k^{(s)} := e_k^{(s)} - \mathbb{E}(e_k^{(s)}|\mathcal{F}_{k-1}) \quad ; \quad \tilde{d}_k^{(s)} := d_k - d_k^{(s)}$$

$$M_k^{(s)} := \sum_{i=1}^{k} d_i^{(s)} \quad ; \quad \tilde{M}_k^{(s)} := M_k - M_k^{(s)}$$

$$T_s := 4\sum_{i=1}^{2^s} \mathbb{E}(|d_i|_{\mathcal{X}}^2|\mathcal{F}_{i-1}) \quad ; \quad T_s^{(s)} := \sum_{i=1}^{2^s} \mathbb{E}(|d_i^{(s)}|_{\mathcal{X}}^2|\mathcal{F}_{i-1}) .$$

Notice that, for every $s \geq 1$,

$$T_s^{(s)} \leq T_s . \tag{26}$$

Let $\beta > 0$ be fixed for the moment. Define the events

$$A_s := \{\frac{\max_{1\leq k\leq 2^s} |M_k|_{\mathcal{X}}}{2^{s/2}(L(s))^{1/2}} > \lambda\} \quad ; \quad B_s := \{\frac{\max_{1\leq k\leq 2^s} |M_k^{(s)}|}{2^{s/2}(L(s))^{1/2}} > \lambda/2\}$$

$$C_s := \{\frac{\max_{1\leq k\leq 2^s} |\tilde{M}_k^{(s)}|_{\mathcal{X}}}{2^{s/2}(L(s))^{1/2}} > \lambda/2\} \quad ; \quad D_s := \{\frac{T_s}{2^s} > \beta\lambda^2\} \quad ; \quad E_s := B_s \cap \{\frac{T_s^{(s)}}{2^s} \leq \beta\lambda^2\} .$$

Using (26), we see that $B_s \cap D_s^c \subset E_s$. In particular, we have

$$A_s \subset B_s \cup C_s \quad ; \quad B_s \subset D_s \cup E_s .$$

Hence,

$$\{\sup_{s\geq S} \frac{\max_{1\leq k\leq 2^s} |M_k|_{\mathcal{X}}}{2^{s/2}(L(s))^{1/2}} > \lambda\} = \bigcup_{s\geq S} A_s \subset (\bigcup_{s\geq S} C_s)\bigcup(\bigcup_{s\geq S} D_s)\bigcup(\bigcup_{s\geq S} E_s) .$$

Now, $\bigcup_{s\geq S} D_s = \{\sup_{s\geq S} \frac{T_s}{2^s} > \beta\lambda^2\}$, hence by Hopf maximal inequality (see e.g. [21, Cor. 2.2, p. 8], using that $\mathbb{E}(|d_1|_{\mathcal{X}}^2) = 1$,

$$\mathbb{P}(\bigcup_{s\geq S} D_s) \leq \mathbb{P}(\bigcup_{s\geq 1} D_s) \leq \frac{4}{\beta\lambda^2} . \tag{27}$$

We also easily see that, using Fubini for the last estimate,

$$\mathbb{P}(\bigcup_{s\geq S} C_s) \leq \frac{2}{\lambda}\sum_{s\geq 0} \frac{\mathbb{E}(\max_{1\leq k\leq 2^s} |\tilde{M}_k^{(s)}|_{\mathcal{X}})}{2^{s/2}(L(s))^{1/2}}$$

$$\leq \frac{4}{\lambda}\sum_{s\geq 1} \frac{2^{s/2}}{(L(s))^{1/2}} \mathbb{E}(|d_1|_{\mathcal{X}} \mathbf{1}_{\{|d_1|_{\mathcal{X}} \geq \alpha\lambda 2^{s/2}/(L(s))^{1/2}\}}) \leq \frac{4C}{\alpha\lambda^2} . \tag{28}$$

It remains to deal with $\bigcup_{r\geq R} E_r$. We need the following lemma from Dedecker-Gouëzel-Merlevède [11, Proposition A.1], whose proof follows from Pinelis [32, Theorem 3.4]. The proof in [11] is done in the scalar case but it easily extends to 2-smooth Banach spaces, since Theorem 3.4 in [32] is proved in that setting. A related inequality in the scalar case is stated in Freedman [18, Theorem 1.6].

Lemma 4.1. *Let $c > 0$. Let $(\mathcal{F}_j)_{j\geq 0}$ be a non-decreasing filtration and $(d_j)_{j\geq 1}$ a sequence of random variables adapted to $(\mathcal{F}_j)_{j\geq 0}$, such that for every $j \geq 1$, $|d_j|_{\mathcal{X}} \leq c$ a.s. and $\mathbb{E}(d_j|\mathcal{F}_{j-1}) = 0$ a.s. Then, for all $x, y > 0$ and all integer $n \geq 1$, we have*

$$\mathbb{P}(\max_{1\leq k\leq n} |\sum_{i=1}^{k} d_i|_{\mathcal{X}} > x \,;\, \sum_{i=1}^{n} \mathbb{E}(|d_i|_{\mathcal{X}}^2|\mathcal{F}_{i-1}) \leq y/D^2) \leq 2\exp\big(-\frac{y}{c^2}h(\frac{xc}{y})\big) , \tag{29}$$

where $h(u) = (1+u)\log(1+u) - u$.

Let $s \geq S$. Let us apply the lemma to the sequence of martingale differences $(d_i^{(s)})$ (in this case, we may take $c = 2\alpha\lambda 2^{s/2}/(L(s))^{1/2})$, with $x = \lambda 2^{s/2-1}(L(s))^{1/2}$, $y = \beta D^2\lambda^2 2^s$ and $n = 2^s$. We obtain, taking $\alpha = D^2\beta$,

$$\mathbb{P}(E_s) \leq 2\exp\big(-\frac{D^2\beta L(s)}{4\alpha^2}h(\frac{\alpha}{D^2\beta})\big) = 2\exp\big(-\frac{L(s)h(1)}{4D^2\beta}\big) = \frac{2}{s^{h(1)/4D^2\beta}}\,.$$

Hence, if $h(1)/(4D^2\beta) > 1$, we see that

$$\sum_{s\geq S}\mathbb{P}(E_s) \leq \frac{2}{(h(1)/4D^2\beta - 1)S^{h(1)/4D^2\beta-1}}\,. \tag{30}$$

Take $\beta = \frac{(2-p)h(1)}{8D^2}$ and $S = [\lambda^{2-p}]$. Recall that we assume that $\lambda \geq D$, in particular $\frac{1}{\lambda^2} \leq \frac{D^{p-2}}{\lambda^p}$. Combining (25), (27), (28) and (30), we infer that, there exists $C > 0$, such that

$$\lambda^p\mathbb{P}(M^* > \lambda) \leq \frac{CD^p}{2-p}\,,$$

which ends the proof of (8).

Let us prove that $(d_n)_{n\geq 1}$ satisfies the CLIL, we use the Banach principle, see Proposition D.1. By definition of the Bochner spaces, there exists $(d^{(m)})_{m\geq 1}$, converging in $L^2(\Omega,\mathcal{X})$ to d, such that for every $m \geq 1$, there exist $k_m \geq 1$, $\alpha_1,\ldots,\alpha_{k_m} \in \mathcal{X}$ and $A_1,\ldots,A_{k_m} \in \mathcal{F}_1$ such that

$$d^{(m)} = \sum_{i=1}^{k_m}\alpha_i\mathbf{1}_{A_i}$$

Write $\tilde{d}^{(m)} := d^{(m)} - \mathbb{E}_0(d^{(m)})$. Then, $(\tilde{d}^{(m)})_{m\geq 1}$ converges in $L^2(\Omega,\mathcal{X})$ to d and for every $m \geq 1$, by the bounded law of the iterated logarithm for martingales with stationary increments, taking values in a *finite* fimensional Banach space (which follows for instance from (8)), we see that for every $m \geq 1$, $(d_n^{(m)})_{n\geq 1}$ satisfies the CLIL. Hence we conclude thanks to Proposition D.1.

It remains to prove (9). By the compact LIL and ergodicity, there exists $S \geq 0$, such that $\limsup \frac{|S_n(d)|_{\mathcal{X}}}{\sqrt{2nL(L(n))}} = S$ $\mathbb{P}$-a.s. Let $M := \sup_{|x^*|_{\mathcal{X}^*}\leq 1}\|x^*(d)\|_2$. Let us prove that $S = M$. Let $\varepsilon > 0$. There exists $x^*_\varepsilon \in \mathcal{X}^*$ such that $\|x^*_\varepsilon(X)\|_2 \geq M - \varepsilon$. Since, $|S_n(d)|_{\mathcal{X}} \geq |x^*_\varepsilon(S_n(d))|$, it follows from the LIL for real-valued martingales (with stationary ergodic increments), that

$$S \geq M - \varepsilon \qquad \mathbb{P}\text{-a.s.}$$

Letting $\varepsilon \to 0$, along rational numbers, we see that $S \geq M$. Let us prove the converse inequality.

By the compact LIL and ergodicity, there exists a compact set $K \in \mathcal{X}$, such that for $\mathbb{P}$-a.e. $\omega \in \Omega$, the cluster set of $\{S_n(d)(\omega)/\sqrt{2nL(L(n))},\ n \geq 1\}$ is K. Let $x \in K$ be such that $|x|_{\mathcal{X}} = S$, and let $x^* \in \mathcal{X}^*$ be such that $|x^*|_{\mathcal{X}^*} = 1$ and $x^*(x) = |x|_{\mathcal{X}}$. For $\mathbb{P}$-a.e. $\omega \in \Omega$, there exists $(n_k = n_k(\omega))_{k\geq 1}$ such that $S_{n_k}(d)(\omega)\sqrt{2n_kL(L(n_k))} \xrightarrow[k\to\infty]{|\cdot|_{\mathcal{X}}} x$. In particular

$$x^*(S_{n_k}(d)(\omega)\sqrt{2n_kL(L(n_k))}) \xrightarrow[k\to\infty]{|\cdot|_{\mathcal{X}}} x^*(x) = S \leq \limsup S_n(x^*(d))(\omega)\sqrt{2nL(L(n))})\,.$$

But, by the real LIL, for $\mathbb{P}$-a.e. $\omega \in \Omega$,

$$\limsup S_n(x^*(d))(\omega)\sqrt{2nL(L(n))}) \leq \|x^*(d)\|_2 \leq M\,,$$

which ends the proof. $\square$

Proof of Remark 2.2b. Classically, it suffices to prove that $\int_2^\infty \frac{t}{L(t)^{2+\varepsilon}}\mathbb{P}(M^*>t)dt<\infty$. Now, combining (25), (27), (28) and (30), we infer that, there exists $C>0$, such that, for every $\lambda>0$ and every $1\le p<2$,

$$\mathbb{P}(M^*>\lambda)\le C\Big(\frac{1}{(2-p)L(\lambda)\lambda^p}+\frac{1}{(2-p)\lambda^2}+\frac{(2-p)}{\lambda^p}\Big)\,.$$

For every $n\ge 1$, write $p_n=2-1/n$. It follows that

$$\begin{aligned}\int_2^\infty \frac{t}{L(t)^{2+\varepsilon}}\mathbb{P}(M^*>t)dt &\le \sum_{n\ge 1}\int_{2^n}^{2^{n+1}}\Big(\frac{t}{L(t)^{2+\varepsilon}}\big(\frac{n}{L(t)t^{p_n}}+\frac{n}{t^2}+\frac{1}{nt^{p_n}}\big)\Big)dt\\ &\le C\sum_{n\ge 1}\big(\frac{1}{n^{2+\varepsilon}}+\frac{1}{n^{1+\varepsilon}}+\frac{1}{n^{3+\varepsilon}}\big)<\infty\,.\end{aligned}$$

5. Proof of the results for stationary processes

5.1. **Proof of Theorem 2.6.** Recall that we assume here θ to be invertible. Let $\mathcal{X}$ be a 2-smooth Banach space.

Define

$$H_2:=\{Z\in L^2(\Omega,\mathcal{X})\ :\mathbb{E}_{-\infty}(Z)=0,\,\mathbb{E}_\infty(Z)=Z,\,\sum_{n\in\mathbb{Z}}\|P_nZ\|_{2,\mathcal{X}}<\infty\}\,. \tag{31}$$

It is not difficult to see that, setting $\|Z\|_{H_2}:=\sum_{n\in\mathbb{Z}}\|P_nZ\|_{2,\mathcal{X}}$, $(H_2,\|\cdot\|_{H_2})$ is a Banach space.

By our regularity conditions, we have, $Z=\sum_{k\in\mathbb{Z}}P_kZ$ in $L^2(\Omega,\mathcal{X})$ and $\mathbb{P}$-a.s. Hence, writing $S_n=S_n(Z)=\sum_{i=0}^{n-1}Z\circ\theta^i$, we have

$$S_n=\sum_{k\in\mathbb{Z}}\sum_{i=0}^{n-1}(P_kZ)\circ\theta^i\,.$$

This splitting of S_n into a series of martingales with (stationary) increments has been used already in [38] and [8] in a similar context. This idea seems to appear first (explicitly) in a paper by McLeish [26]. We deduce that

$$\mathcal{M}_2(Z)\le\sum_{k\in\mathbb{Z}}\mathcal{M}_2(P_k(Z))\,.$$

But, for every $k\in\mathbb{Z}$, $((P_kZ)\circ\theta^i))_{i\ge 1}$ is a stationary sequence of martingale differences. Hence, by Theorem 2.2, for every $1\le p<2$, there exists C_p, such that

$$\|\mathcal{M}_2(Z)\|_{p,\infty}\le C_pD(\sum_{k\in\mathbb{Z}}\|P_kZ\|_2)\,. \tag{32}$$

We define a continuous operator $\mathcal{D}$ from H_2 with values in $\{d\in L^2(\Omega,\mathcal{F}_1)\ :\ \mathbb{E}(d_1|\mathcal{F}_0)=0\}$, by setting, for every $Z\in H_2$, $\mathcal{D}Z:=\sum_{n\in\mathbb{Z}}P_1(Z\circ\theta^n)$. Write $d=\mathcal{D}Z$. Let $M_n:=\sum_{i=0}^{n-1}d\circ\theta^k$. We want to prove that

$$|S_n-M_n|_{\mathcal{X}}=o(\sqrt{nL(L(n))})\qquad\mathbb{P}\text{-a.s.}\,. \tag{33}$$

Since $\mathcal{M}_2(Z+d)\le\mathcal{M}_2(Z)+\mathcal{M}_2(d)$, using (32), Theorem 2.2 and the Banach principle (see the appendix), we see that the set of $Z\in H_2$ such that (33) holds is closed in H_2.

Let $H_2^-:=\{Z\in H_2\ :\ Z\in L^2(\Omega,\mathcal{F}_0)\}$ and $H_2^+:=\{Z\in H_2\ :\ \mathbb{E}_0(Z)=0\}$. We have the direct sum $H_2=H_2^-\oplus H_2^+$. We prove the result on each space separately.

We define two operators Q and R, acting respectively on $\{Z\in L^1(\Omega,\mathcal{F}_0)\ :\ \mathbb{E}_{-\infty}(Z)=0\}$ and on $\{Y\in L^1(\Omega,\mathcal{F}_\infty)\ :\ \mathbb{E}_0(Y)=0\}$, by setting

$$QZ=\mathbb{E}_0(Z\circ\theta)\ \text{ and }\ RY=Y\circ\theta^{-1}-\mathbb{E}_0(Y\circ\theta^{-1})\,.$$

Those operators have been already used in [35] (see also [8] and [6] where Q has been used in a similar context). Notice that for every $n \geq 1$,

$$Q^n Z = \mathbb{E}_0(Z \circ \theta^n) \text{ and } R^n Y = Y \circ \theta^{-n} - \mathbb{E}_0(Y \circ \theta^{-n}).$$

In particular, for every $Z \in H_2^-$,

$$\|Q^n Z\|_{H_2^-} = \sum_{k \geq 0} \|P_{-k}(\mathbb{E}_0(Z \circ \theta^n))\|_2 = \sum_{k \geq n} \|P_{-k} Z\|_2 \underset{n \to \infty}{\longrightarrow} 0$$

and, for every $Y \in H_2^+$,

$$\|R^n Y\|_{H_2^+} = \sum_{k \geq 0} \|P_k(Y \circ \theta^{-n} - \mathbb{E}_0(Y \circ \theta^{-n}))\|_2 = \sum_{k \geq n} \|P_k Y\|_2 \underset{n \to \infty}{\longrightarrow} 0$$

This implies that Q and R are contractions of H_2^- and H_2^+ respectively, and that they satisfy the mean ergodic theorem, see e.g. Theorems 1.2 and 1.3 page 73 of [21]. In particular, we have

$$H_2^- = \overline{(I-Q)H_2^-}^{H_2} \text{ and } H_2^+ = \overline{(I-R)H_2^+}^{H_2}.$$

Hence, we just have to prove that (33) holds on $(I-Q)H_2^-$ and $(I-R)H_2^+$. Those cases actually "correspond" to a martingale-coboundary decomposition as in Gordin-Lifšic [19]. Indeed, for every $n \in \mathbb{Z}$ and every $Z \in H_2^-$, we have

$$P_1\big(((I-Q)Z) \circ \theta^n\big) = (P_{1-n} Z) \circ \theta^n - \big(P_{-n}(\mathbb{E}_{-1}(Z))\big) \circ \theta^{n+1}.$$

Hence, since $Z \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$, we have

$$\mathcal{D}((I-Q)Z) = \sum_{n \geq 1} \big((P_{-(n-1)} Z) \circ \theta^n - (P_{-n} Z) \circ \theta^{n+1}\big) = (P_0 Z) \circ \theta,$$

using telescopic sums and the fact that $\|P_{-n} Z\|_2 \to 0$, as $n \to \infty$. Then, we have the decomposition,

$$(I-Q)Z - \mathcal{D}((I-Q)Z) = Z - Z \circ \theta,$$

which implies that $|S_n((I-Q)Z) - M_n((I-Q)Z)|_{\mathcal{X}} = o(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s., since, by the Borel-Cantelli lemma, $Z \circ \theta^n / \sqrt{n} \to 0$ $\mathbb{P}$-a.s., as $n \to \infty$.

Let us prove (33) on $(I-R)H_2^+$. For every $Y \in H_2^+$, we have

$$P_1\big(((I-R)Y) \circ \theta^n\big) = P_1\big(Y \circ \theta^n - Y \circ \theta^{n-1}\big) + \big(P_{2-n}(\mathbb{E}_1(Y))\big) \circ \theta^{n-1}.$$

Hence, since $\mathbb{E}_0(Y) = 0$,

$$\mathcal{D}((I-Q)Y) = \mathbb{E}_1(Y) = P_0(Y).$$

Then, we have the decomposition,

$$(I-R)Y - \mathcal{D}((I-R)Y) = (Z - \mathbb{E}_1(Y)) - (Y - \mathbb{E}_1(Y)) \circ \theta^{-1},$$

which implies that $|S_n((I-R)Y) - M_n((I-R)Y)|_{\mathcal{X}} = o(\sqrt{nL(L(n))})$ $\mathbb{P}$-a.s., since, by the Borel-Cantelli lemma, $(Y - \mathbb{E}_1(Y)) \circ \theta^{n-1} / \sqrt{n} \to 0$ $\mathbb{P}$-a.s., as $n \to \infty$. $\square$

5.2. **Proof of Remark 2.6a.** By (12), we have $\sum_{n \in \mathbb{Z}} \|P_1(X_n)\|_{2,\mathcal{X}} < \infty$. Hence, for every $f, g \in \mathcal{X}^*$, we have, with absolute convergence of all the series,

$$\begin{aligned}\mathcal{K}_d(f,g) &= \sum_{m,n \in \mathbb{Z}} \mathbb{E}(P_1(f(X_n)) P_1(g(X_m))) = \sum_{m,n \in \mathbb{Z}} \mathbb{E}(f(X_0) P_{1-n}(g(X_{m-n}))) \\ &= \sum_{m,n \in \mathbb{Z}} \mathbb{E}(f(X_0) P_{-n}(g(X_m))) = \sum_{m \in \mathbb{Z}} \mathbb{E}(f(X_0) g(X_m)).\end{aligned}$$

5.3. **Proof of Theorem 2.5.** As in the proof of Theorem 2.6, we define a Banach space

$$H_p := \{Z \in L^p(\Omega, \mathcal{X}) \ : \mathbb{E}_{-\infty}(Z) = 0,\, \mathbb{E}_{\infty}(Z) = Z,\, \|Z\|_{H_p} := \sum_{n\in\mathbb{Z}} \|P_n Z\|_{p,\mathcal{X}} < \infty\} .$$

We see that

$$\|\mathcal{M}_2 Z\|_{p,\infty} \le C_{p,r} D^{1/p} \|Z\|_{H_p} ,$$

where $C_{r,p}$ is the constant appearing in Proposition 2.1, and that the operator $\mathcal{D}$ may be extended in a continuous operator from H_p to $\{d \in L^p(\Omega, \mathcal{F}_1, \mathcal{X}) \ : \ \mathbb{E}_0(d) = 0\}$. Then, the proof is the same, using that the Marcinkiewicz-Zygmund strong law of large number is known for r-smooth valued stationary martingale differences, see e.g. [37]. $\square$

5.4. **Proof of Theorem 2.7.** For every $n \ge 0$ define $P^{(n)} := \mathbb{E}^n - \mathbb{E}^{n+1}$. It suffices to prove the theorem under the weaker condition $\mathbb{E}^\infty(X) = 0$ and

$$\sum_{n\ge 0} \|P^{(n)}(X)\|_{2,\mathcal{H}} < \infty .$$

The fact that (20) implies the latter may be proved as Remark 2.6b, using (19).

Then, the proof may be done exactly as the proof of Theorem 2.6 on H_2^-, the role of Q being played by $\mathbf{K}$.

Appendix A. Proof of Proposition 2.1

We start with the case $d \in L^p(\Omega, \mathcal{F}_1, \mathbb{P})$ and $\mathbb{E}_0(d) = 0$. Define $M^* = M^*(d) := \sup_{s\ge 0} \frac{\max_{1\le n\le 2^s} |S_n(d)|_{\mathcal{X}}}{2^{s/p}}$. Let $s \ge 0$. For every $2^s \le n \le 2^{s+1} - 1$, we have

$$\frac{|S_n(d)|_{\mathcal{X}}}{n^{1/p}} \le \frac{\max_{1\le n\le 2^s} |S_n(d)|_{\mathcal{X}}}{2^{s/p}} \le M^* .$$

Hence it suffices to prove the result for M^* instead of $\mathcal{M}_p(d)$. Let $\lambda > 0$. We proceed by truncation. For every $s \ge 0$, $k \ge 1$ define

$$e_k^{(s)} := d_k \mathbf{1}_{\{|d_k|_{\mathcal{X}} \le \lambda 2^{s/p}\}} \quad ; \quad d_k^{(s)} := e_k^{(s)} - \mathbb{E}(e_k^{(s)}|\mathcal{F}_{k-1}) \quad ; \quad \tilde{e}_k^{(s)} := d_k - e_k^{(s)} \quad ;$$
$$\tilde{d}_k^{(s)} := d_k - d_k^{(s)} \quad ; \quad M_k^{(s)} := \sum_{i=1}^{k} d_i^{(s)} \quad ; \quad \tilde{M}_k^{(s)} := M_k - M_k^{(s)} .$$

Let $\lambda > 0$. Then,

$$\begin{aligned} &\mathbb{P}(M^* > \lambda) \\ &\le \sum_{s\ge 0} \mathbb{P}\Big(\frac{\max_{1\le n\le 2^s} |\tilde{M}_n^{(s)}|_{\mathcal{X}}}{2^{s/p}} > \lambda/2\Big) + \sum_{s\ge 0} \mathbb{P}\Big(\frac{\max_{1\le n\le 2^s} |M_n^{(s)}|_{\mathcal{X}}}{2^{s/p}} > \lambda/2\Big) \\ &\le \frac{4}{\lambda} \sum_{s\ge 0} 2^{(1-1/p)s} \mathbb{E}(|\tilde{e}_1^{(s)}|_{\mathcal{X}}) + \frac{2^r}{\lambda^r} \sum_{s\ge 0} \frac{\mathbb{E}(\max_{1\le n\le 2^s} |M_n^{(s)}|_{\mathcal{X}}^r)}{2^{rs/p}} . \end{aligned}$$

Now, by Fubini and stationarity,

$$\sum_{s\ge 0} 2^{(1-1/p)s} \mathbb{E}(|\tilde{e}_1^{(s)}|_{\mathcal{X}}) \le \frac{C\mathbb{E}(|d_1|_{\mathcal{X}}^p)}{\lambda^p}.$$

To deal with the second term, we use Doob's maximal inequality in L^r, for the submartingale $(|M_n|_{\mathcal{X}})_{n\geq 1}$, and (2). We obtain

$$\sum_{s\geq 0}\frac{\mathbb{E}(\max_{1\leq n\leq 2^s}|M_n^{(s)}|_{\mathcal{X}}^r)}{2^{rs/p}}\leq \sum_{s\geq 0}\frac{C_r}{2^{rs/p}\lambda^r}\mathbb{E}(|M_{2^s}^{(s)}|_{\mathcal{X}}^r) \tag{34}$$
$$\leq D^rC_r\sum_{s\geq 0}2^{(1-r/p)s}\mathbb{E}(|d_1^{(s)}|_{\mathcal{X}}^r)\leq \frac{D^rC_{r,p}\mathbb{E}(|d_1|_{\mathcal{X}}^p)}{\lambda^{p-r}}\,,$$

which proves the proposition, in that case. When $d\in L^2(\Omega,\mathcal{F}^0,\mathbb{P})$ and $\mathbb{E}^1(d)=0$, the proof is the same, with the obvious changes, noticing that for every $n\geq 1$, $(S_n(d)-S_{n-k}(d))_{0\leq k\leq n}$ is a $(\mathcal{F}^{n-k})_{0\leq k\leq n}$-martingale and that $\max_{1\leq k\leq n}|S_k(d)|_{\mathcal{X}}\leq 2\max_{1\leq k\leq n}|S_n(d)-S_{n-k}(d)|_{\mathcal{X}}$ □

Appendix B. Proof of Remarks 2.6b and 2.5

We assume that $\mathcal{X}$ is 2-convex For every $n\geq 0$, using Cauchy-Schwarz and (**??**), we have

$$\Big(\sum_{k=2^n}^{2^{n+1}-1}\|P_{-k}X\|_{2,\mathcal{X}}\Big)^2\leq 2^n\sum_{k\geq 2^n}\mathbb{E}(|P_{-k}X|_{\mathcal{H}}^2)\leq C2^n\mathbb{E}(|\mathbb{E}_{-2^n}(X)|_{\mathcal{X}}^2)\,,$$
$$\text{and }\Big(\sum_{k=2^n}^{2^{n+1}-1}\|P_kX\|_{2,\mathcal{X}^2}\Big)^2\leq 2^n\sum_{k\geq 2^n}\mathbb{E}(|P_kX|_{\mathcal{H}}^2)\leq C2^n\mathbb{E}(|X-\mathbb{E}_{2^n}(X)|_{\mathcal{X}}^2)\,,$$

and Remark 2.6b follows, since the sequences $(\|\mathbb{E}_{-n}(X)\|_{2,\mathcal{H}})$ and $(\|X-\mathbb{E}_n(X)\|_{2,\mathcal{H}})$ are non-increasing.

Let us prove Remark 2.5. For every $1<p<2$, by Hölder's inequality twice we have, with $1/p+1/q=1$,

$$\Big(\sum_{k=2^n}^{2^{n+1}-1}\|P_{-k}X\|_{p,\mathcal{H}}\Big)^p\leq 2^{np/q}\mathbb{E}\Big(\sum_{k=2^n}^{2^{n+1}-1}|P_{-k}X|_{\mathcal{H}}^p\Big)\leq 2^{np/2}\mathbb{E}\Big(\big(\sum_{k\geq 2^n}|P_{-k}X|_{\mathcal{H}}^2\big)^{p/2}\Big)\leq C2^{np/2}\|\mathbb{E}_{-2^n}(X)\|_{p,\mathcal{H}}^p\,,$$
$$\text{and }\Big(\sum_{k=2^n}^{2^{n+1}-1}\|P_kX\|_{p,\mathcal{H}}\Big)^2\leq 2^{np/q}\mathbb{E}\Big(\sum_{k=2^n}^{2^{n+1}-1}|P_kX|_{\mathcal{H}}^p\Big)\leq 2^{np/2}\mathbb{E}\Big(\big(\sum_{k\geq 2^n}|P_kX|_{\mathcal{H}}^2\big)^{2/p}\Big)\leq C2^{np/2}\|X-\mathbb{E}_{2^n}(X)\|_{p,}^p$$

where we used Burkholder's inequality in Hilbert spaces, see [3]. Then, we conclude as above.

Appendix C. Proof of Corollary 3.3

Notice that, by (23), for every $x,h,h'\in\mathbb{R}$, we have

$$|f(x+h)-f(x+h')|\leq 2^r\varphi(|h-h'|)(1+|x|^r)+2^{r-1}K(|h|^r+|h'|^r)\,. \tag{35}$$

Recall that for every concave ψ with $\psi(0)=0$, $x\to\psi(x)/x$ is non-increasing on $]0,+\infty[$ and ψ is sub-additive.

We want to apply Theorem 2.6 and Remark 2.6b. We shall evaluate $\|P_0(X_n)\|_2$, $\|\mathbb{E}_0(X_n)\|_2$ and $\|X_n-\mathbb{E}_n(X_n)\|$.

Enlarging our probability space if necessary, we assume that there exists (ξ_n') an independent copy of (ξ_n).

Then,

$$P_0X_n=\mathbb{E}_0\big(f(A_n+h_n)-f(A_n+h_n')\big)\,,$$

where $A_n:=\sum_{k>-n}a_{-k}\xi_{n+k}'+\sum_{k>n}a_k\xi_{n-k}$, $h_n:=a_n\xi_0$ and $h_n':=a_n\xi_0'$.

In particular, we have, by independence and using (35),

$$\mathbb{E}((P_0X_n)^2)\leq C_r\big(\mathbb{E}(\varphi^2(|a_n|(|\xi_0|+|\xi_0'|)\,\mathbb{E}(|A_n|^{2r})+|a_n|^{2r}\mathbb{E}(|\xi_0|^{2r}\,)\big)$$

We notice now that for every $\varphi \in \Lambda$, there exists $C > 0$ such that, for every $n \geq 1$

$$\mathbb{E}(\varphi^2(|a_n|(|\xi_0| + |\xi_0'|)) \leq C\varphi^2(|a_n|)\,. \tag{36}$$

This follows from Jensen's inequality and the sub-additivity of φ^2 (using that $\xi_0 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$), and it is obvious when $\varphi(x) = \min(1, x^\alpha)$ (using that $\xi_0 \in L^{2\alpha}(\Omega, \mathcal{F}, \mathbb{P})$).

Clearly, $\mathbb{E}(|A_n|^{2r}) \leq \big(\sum_{k\in\mathbb{Z}} |a_k| \|\xi_0\|_{2r}\big)^{2r}$.

Since $x \to \varphi^2(x)/x$ is non-decreasing, when φ^2 is concave, we see that whenever $\varphi \in \Lambda$, $|a_n|^{2r} \leq C\varphi^2(|a_n|)$.

This finishes the proof of Corollary 3.3 under the assumption onn $P_0(X_n)$.

We shall now evaluate $\|\mathbb{E}_0(X_n)\|_2$ and $\|X_n - \mathbb{E}_n(X_n)\|_2$. We have

$$\mathbb{E}_0(X_n) = \mathbb{E}_0(f(B_n + k_n) - f(B_n - k_n'))\,,$$

where $B_n := \sum_{k>-n} a_{-k}\xi_{n+k}$, $k_n = \sum_{k\geq n} a_k \xi_{n-k}$ and $k_n' = \sum_{k\geq n} a_k \xi_{n-k}'$. Hence, using (35),

$$\|E_0(X_n)\|_2^2 \leq C_r\big(\mathbb{E}(\varphi^2(|k_n| + |k_n'|)\,\mathbb{E}(|A_n|^{2r}) + 2\|k_n\|_{2r}^{2r}\,)\big)\,.$$

When φ^2 is concave, by Jensen's inequality,

$$\mathbb{E}(\varphi^2(|k_n| + |k_n'|) \leq \varphi^2(2\mathbb{E}(|\xi_0|)\sum_{k\geq n} |a_k|) \leq (1 + 2\mathbb{E}(|\xi_0|))\varphi^2(\sum_{k\geq n} |a_k|)$$

When $\varphi(x) = \min(1, x^\alpha)$, assuming that $1/2 \leq \alpha \leq 1$ (otherwise we are in the previous case), we have

$$\mathbb{E}(\varphi^2(|k_n| + |k_n'|) \leq \big(\sum_{k\geq n} |a_k| \|\xi_0\|_{2\alpha}\big)^{2\alpha} \leq C\varphi^2(\sum_{k\geq n} |a_k|)$$

Clearly, $\mathbb{E}(|B_n|^{2r}) \leq \big(\sum_{k\in\mathbb{Z}} |a_k| \|\xi_0\|_{2r}\big)^{2r}$.

Finally, we have

$$\|k_n\|_{2r}^{2r} \leq \|\xi_0\|_{2r}^{2r}\big(\sum_{k\geq n} |a_k|\big)^{2r}\,.$$

Since $x \to \varphi^2(x)/x$ is non-decreasing, when φ^2 is concave, we see that whenever $\varphi \in \Lambda$,

$$\|k_n\|_{2r}^{2r} \leq C\varphi^2(\sum_{k\geq n} |a_k|)\,.$$

Appendix D. The Banach principle

The following is an extension of the Banach principle as stated in Theorem 7.2 p. 64 of [21].

Proposition D.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{X}, \mathbf{B}$ be Banach spaces. Let $\mathcal{C}$ be a vector space of measurable functions from Ω to $\mathcal{X}$. Let $(T_n)_{n\geq 1}$ be a sequence of linear maps from $\mathbf{B}$ to $\mathcal{C}$. Assume that there exists a positive decreasing function L on $]0, +\infty[$, with $\lim_{\lambda\to\infty} L(\lambda) = 0$, such that*

$$\mathbb{P}(\sup_{n\geq 1} |T_n x|_{\mathcal{X}} > \lambda |x|_{\mathbf{B}}) \leq L(\lambda) \qquad \forall \lambda > 0, x \in \mathbf{B}\,. \tag{37}$$

Then the sets $\{x \in \mathbf{B} : |T_n x|_{\mathcal{X}} \to 0 \quad \mathbb{P}\text{-a.s.}\}$ and $\{x \in \mathbf{B} : (T_n x)_{n\geq 1}$ is $\mathbb{P}$-a.s. relatively compact in $\mathcal{X}$ are closed in $\mathbf{B}$.

Proof. We prove that the second set is closed, the proof for the first one being similar, but easier. Let $x \in \mathbf{B}$ and $(x_m)_{m\geq 1} \subset \mathbf{B}$ be such that $|x_m - x|_{\mathbf{B}} \underset{m\to\infty}{\longrightarrow} 0$ and such that for every $m \geq 1$, $(T_n x_m)_{n\geq 1}$ is $\mathbb{P}$-a.s. relatively compact in $\mathcal{X}$. We want to prove that $(T_n x)_{n\geq 1}$ is $\mathbb{P}$-a.s. relatively compact.

By (37), for every integers $m, p \geq 1$,

$$\mathbb{P}(\sup_{n\geq 1} |T_n(x - x_m)|_{\mathcal{X}} > 1/p) \leq L\big(\frac{1}{p|x - x_m|_{\mathbf{B}}}\big) \qquad \forall \lambda > 0, x \in \mathbf{B}\,.$$

Since $\lim_{\lambda\to\infty} L(\lambda) = 0$, there exists a subsequence $(m_k)_{k\geq 1}$ and a set $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$, such that for every $\omega \in \Omega_0$,

$$\sup_{n\geq 1} |T_n(x - x_{m_k})|_{\mathcal{X}}(\omega) \underset{k\to\infty}{\longrightarrow} 0 .$$

There exists $\Omega_1 \in \mathcal{F}$, with $\mathbb{P}(\Omega_1) = 1$, such that, for every $\omega \in \Omega_1$ and every $k \geq 1$, $((T_n x_{m_k})(\omega))_{n\geq 1}$ is relatively compact in $\mathcal{X}$.

Let $\omega \in \Omega_0 \cap \Omega_1$ be fixed. Let φ_0 be an increasing function from $\mathbb{N}$ to $\mathbb{N}$. We want to prove that $(T_{\varphi_0(n)} x(\omega))_{n\geq 1}$ admits a convergent subsequence.

For every $k \geq 1$, $((T_{\varphi_0(n)} x_{m_k})(\omega))_{n\geq 1}$ admits a Cauchy subsequence. We construct by induction some increasing functions $(\varphi_k)_{k\geq 1}$ such that, for every $k \geq 1$, setting $\psi_k := \varphi_0 \circ \varphi_1 \circ \cdots \circ \varphi_k$, we have for every $p \geq n \geq 1$,

$$|T_{\psi_k(n)} x_{m_k}(\omega) - T_{\psi_k(p)} x_{m_k}(\omega)|_{\mathcal{X}} \leq 1/n .$$

Then, $(T_{\psi_n(n)} x(\omega))$ is Cauchy. Indeed, for every $N \geq 1$, and every $p > n \geq N$, we have

$$\begin{aligned} &|T_{\psi_n(n)} x(\omega) - T_{\psi_p(p)} x(\omega)|_{\mathcal{X}} \\ &\leq |T_{\psi_n(n)} x_{m_n}(\omega) - T_{(\psi_n\circ\varphi_{n+1}\circ\cdots\circ\varphi_p)(p))} x_{m_n}(\omega)|_{\mathcal{X}} + 2\sup_{r\geq 1} |T_r(x_{m_n} - x)|_{\mathcal{X}} \underset{N\to\infty}{\longrightarrow} 0 , \end{aligned}$$

and the result follows. □

References

[1] G. Argiris and J. Rosenblatt, *Forcing divergence when the supremum is not integrable*, Positivity 10 (2006), no. 2, 261-284.

[2] E. Berger, *An almost sure invariance principle for stationary ergodic sequences of Banach space valued random variables*, Probab. Theory Related Fields 84 (1990), no. 2, 161-201.

[3] D. L. Burkholder, *Sharp inequalities for martingales and stochastic integrals. Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987).* Astérisque No. 157-158 (1988), 75-94.

[4] C. Cuny, *On the forward-backward martingale decomposition*, unpublished notes.

[5] C. Cuny, *Pointwise ergodic theorems with rate with applications to limit theorems for stationary processes*, Stoch. Dyn. 11 (2011), no. 1, 135155

[6] C. Cuny. and F. Merlevède, *On martingale approximations and the quenched weak invariance principle.* arXiv:1202.2964

[7] C. Cuny and F. Merlevède, *Strong invariance principles with rate for "reverse" martingales and applications*, submitted.

[8] C. Cuny and D. Volný, *A quenched invariance principle for stationary processes*, submitted.

[9] J. A. Davis, *Convergence rates for the law of the iterated logarithm*, Ann. Math. Statist. 39 1968 1479?1485.

[10] J. Dedecker, P. Doukhan and F. Merlevède, *Rates of convergence in the strong invariance principle under projective criteria*, arXiv:1102.5061.

[11] J. Dedecker, S. Gouëzel and F. Merlevède, *Some almost sure results for unbounded functions of intermittent maps and their associated Markov chains*, Ann. Inst. Henri Poincaré Probab. Stat. 46 (2010), no. 3, 796-821.

[12] J. Dedecker, S. Gouëzel and F. Merlevède, *The almost sure invariance principle for unbounded functions of expanding maps*, (ALEA 9:141-163, 2012)

[13] J. Dedecker and F. Merlevède, *The empirical distribution function for dependent variables: asymptotic and nonasymptotic results in L^p*, ESAIM Probab. and Stat. 11, (2007) 102-114.

[14] J. Dedecker and F. Merlevède, *Convergence rates in the law of large numbers for Banach-valued dependent variables*, Teor. Veroyatn. Primen. 52 (2007), no. 3, 562-587; translation in Theory Probab. Appl. 52 (2008), no. 3, 416–438.

[15] J. Dedecker and F. Merlevède, *On the almost sure invariance principle for stationary sequences of Hilbert-valued random variables*, Dependence in probability, analysis and number theory, 157-175, Kendrick Press, Heber City, UT, 2010.

[16] J. Dedecker, F. Merlevède and F. Pène, *Rates in the strong invariance principle for ergodic automorphisms of the torus*, arXiv:1206.4336.

[17] H. Dehling, M. Denker and W. Philipp, *A bounded law of the iterated logarithm for Hilbert space valued martingales and its application to U-statistics*, Probab. Theory Relat. Fields 72 (1986), no. 1, 111-131.

[18] D. A. Freedman, *On tail probabilities for martingales*, Ann. Probab. 3 (1975), 100-118.

[19] M. I. Gordin and B. A. Lifšic, *Central limit theorem for stationary Markov processes*, (Russian) Dokl. Akad. Nauk SSSR 239 (1978), no. 4, 766-767.

[20] Y. Jiang and L. Wu, *Hilbertian invariance principle for empirical process associated with a Markov process*, Chinese Ann. Math. Ser. B 24 (2003), no. 1, 1-16.

[21] U. Krengel, Ergodic theorems, de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, (1985).

[22] J. Kuelbs, *A strong convergence theorem for Banach space valued random variables*, Ann. Probab. 4 (1976), no. 5, 744-771.

[23] M. Ledoux and M. Talagrand, *Characterization of the law of the iterated logarithm in Banach spaces*, Ann. Probab. 16 (1988), no. 3, 1242-1264.

[24] Z. Lin and W. Liu, *Strong approximation for a class of stationary processes*, Stochastic Process. Appl. 119 (2009), no. 1, 249-280.

[25] M. Ledoux and M. Talagrand, Probability in Banach spaces. Isoperimetry and processes. Springer-Verlag, Berlin, 1991. xii+480 pp.

[26] D. L. McLeish, *A maximal inequality and dependent strong laws*, Ann. Probability 3 (1975), no. 5, 829?839.

[27] F. Merlevède, *On a maximal inequality for strongly mixing random variables in Hilbert spaces. Application to the compact law of the iterated logarithm*, Ann. I.S.U.P. 52 (2008), no. 1-2, 47?60.

[28] F. Merlevède and M. Peligrad, *Rosenthal inequalities for martingales and stationary sequences and examples.* to appear in Ann. Probab. (2012). arXiv:1103.3242.

[29] F. Merlevède, M. Peligrad and S. Utev, *Recent advances in invariance principles for stationary sequences*, Probab. Surv. 3 (2006), 1-36.

[30] G. J. Morrow and W. Philipp, *An almost sure invariance principle for Hilbert space valued martingales*, Trans. Amer. Math. Soc. 273 (1982), no. 1, 231-251.

[31] W. Philipp, *A note on the almost sure approximation of weakly dependent random variables*, Monatsh. Math. 102 (1986), no. 3, 227-236.

[32] I. Pinelis, *Optimum bounds for the distributions of martingales in Banach spaces*, Ann. Probab. 22 (1994), no. 4, 1679-1706.

[33] G. Pisier, *Martingales with values in uniformly convex spaces*, Israel J. Math. 20 (1975), no. 3-4, 326-350.

[34] D. Volný, *Martingale approximation of non adapted stochastic processes with nonlinear growth of variance*, Dependence in probability and statistics, 141-156, Lecture Notes in Statist., 187, Springer, New York, 2006.

[35] D. Volný, *A nonadapted version of the invariance principle of Peligrad and Utev*, C. R. Math. Acad. Sci. Paris 345 (2007), no. 3, 167-169.

[36] D. Volný and M. Woodroofe, *An example of non-quenched convergence in the conditional central limit theorem for partial sums of a linear process*, Dependence in probability, analysis and number theory, 317?322, Kendrick Press, Heber City, UT, 2010.

[37] W. A., Woyczyński, *Asymptotic behavior of martingales in Banach spaces. II*, Martingale theory in harmonic analysis and Banach spaces (Cleveland, Ohio, 1981), pp. 216-225, Lecture Notes in Math., 939, Springer, Berlin-New York, 1982.

[38] W. B. Wu, *Strong invariance principles for dependent random variables.* Ann. Probab. 35 (2007), 2294–2320.

Laboratoire MAS, Ecole Centrale de Paris, Grande Voie des Vignes, 92295 Chatenay-Malabry cedex, FRANCE.

E-mail address: `christophe.cuny@ecp.fr`